

Methods of information systems synthesis

УДК 004.89

DOI: <https://doi.org/10.20998/2522-9052.2021.4.03>

О. В. Барабаш, О. І. Бандурка, В. В. Шпурик, О. В. Свинчук

Національний технічний університет України «КПІ імені Ігоря Сікорського», Київ, Україна

ІНФОРМАЦІЙНА СИСТЕМА АНАЛІЗУ ГЕОДАНИХ ДЛЯ ВІДСЛІДКОВУВАННЯ ЗМІН РОСЛИННОСТІ

Анотація. Швидке зростання можливостей геоінформаційних технологій в області обробки та аналізу просторових даних призвело до істотного зростання ролі геоінформаційних систем в різних областях людської діяльності. Застосування нових підходів до обробки просторової інформації з супутників для більш ефективної та оперативної оцінки стану рослинних покривів обумовлено зростаючою тенденцією доступності до даних дистанційного зондування Землі. У статті запропоновано інформаційну систему, яка дозволяє швидко та зручно відслідкувати зміни рослинного світу. Аналіз виконувався на прикладі Чорнобильської зони за період 2000-2020 років. Чорнобильська катастрофа співпала з періодом інтенсивного вегетаційного розвитку рослин. У цей період вони є найбільш чутливими до радіації. Було встановлено, що найчастіше для визначення кількісного стану біомаси застосовується індекс NDVI у різні часові інтервали. Але даний індекс стає не ефективним в періоди послаблення активної фази вегетації. У зв'язку з цим представляється практичний інтерес для оцінки можливості застосування методом К-середніх для аналізу космічних знімків рослинного покриву на різних фазах вегетації. В результаті дослідження коректно інтерпретовано водну поверхню, землі з рослинним покривом і без нього, завдяки чому визначено землі з розрідженою рослинністю та щільним рослинним покривом. Побудовано карти рослинного покриву за нормалізованим вегетаційним індексом за методом К-середніх, за яким можна чітко прослідкувати динаміку зміни рослинності протягом 20 років. Точність результатів була перевірена методом Байєса. Згідно проведених розрахунків визначено, що незважаючи на всі природні катаклізми (підвищення температури, засуха, зимові аномалії опадів та температур, буревії, лісові пожежі), а також людську діяльність (санітарні рубки, несанкціонована вирубка дерев), рослинність в Чорнобильській зоні продовжує рости і їхні площі зростатимуть, хоча і не так швидко.

Ключові слова: дистанційне зондування Землі; атмосферна корекція космічних знімків; Quantum GIS; нормалізований диференціальний індекс рослинності; алгоритм К-Means; рослинність.

Вступ

В сучасному інформаційному просторі особливе місце посідають геоінформаційні системи (ГІС), які є унікальним інструментом для створення, управління чи аналізу геопросторових даних в багатьох сферах людської діяльності. Геоінформаційні системи є інтеграцією електронного просторово-орієнтованого зображення (плани, схеми, карти) та бази даних в єдиному інформаційному середовищі [1, 2-5]. Вони є оптимальною платформою для розробки та прийняття системних рішень в сфері екології, в тому числі і для оцінки та аналізу стану рослинності, що зазнали негативного антропогенного та природного впливу. Однією з найбільших загроз, що спіткали Україну, стала аварія на Чорнобильській атомній електростанції (ЧАЕС) із загрозою надзвичайно тяжких наслідків, які ми маємо ще відслідковувати протягом тривалого часу. У зв'язку з цим, аналіз екологічних наслідків поблизу Чорнобиля має важливе наукове значення не лише для України, але і всього світу, і, особливо, сусідніх держав, які теж є постраждалими.

Чорнобильська катастрофа співпала з періодом інтенсивного вегетаційного розвитку рослин. У цей період вони є найбільш чутливими до радіації. У 1986-1988 рр. в 30-кілометровій зоні навколо Чорнобильської АЕС виявлено багато фактів радіаційного пошкодження рослин. Після аварії пошкоджені екосистеми відновлювалися десятиліттями. Таким чином, відбувалось «заселення» рослинних угруповань в колишні міста та села, а також на території сіль-

ськогосподарських угідь. Чорнобильська аварія показала, що великомасштабні викиди радіонуклідів в навколишнє середовище можуть мати серйозні наслідки для біоти, включаючи порушення в екосистемі.

Найбільшого забруднення радіоактивними речовинами зазнали саме природні екосистеми, а особливо ліси, які найважче піддаються стабілізації радіаційної ситуації і потенційно сприяють розповсюдженню радіонуклідів. Наслідком даної тенденції є висихання зелених насаджень на територіях в тисячі гектарів через вплив шкідників, пожежі, хвороби, зміни гідрологічного режиму. Не менш руйнівними та загрозливими є незаконні та самовільні вирубки. Зміна клімату, недбале ставлення і відсутність системи охорони призвели до виникнення неконтрольованих та катастрофічних пожеж [6]. За останнє десятиріччя відмічалися інтенсивні процеси заростання (заліснення) лук та перелогів Чорнобильської зони відчуження деревною рослинністю. Проте в квітні 2020 року було знищено понад 100 га лісу. Пожежа знищила рослинність, яка в Чорнобильській зоні виконує важливу бар'єрну функцію, запобігаючи ерозії ґрунтів.

Використання супутникових спостережень є одним з ефективних методів вивчення трансформації рослинності. Значення індексів рослинності тісно пов'язане з обсягом зеленої біомаси і служить основою для оцінки сезонної і багаторічної динаміки рослин. Таким чином, аналіз часових рядів просторових зображень став основою для створення інформаційної системи аналізу геоданих для відслідковування змін рослинності в Чорнобильській зоні.

Аналіз останніх досліджень і публікацій. Аналізуючи науково-технічні та інші джерела інформації стосовно дослідження стану лісових насаджень на основі використання картографії, геоінформаційних систем, дистанційного зондування Землі (ДЗЗ) та комп'ютерних засобів, варто відзначити декілька методів класифікації космічних знімків: дешифрування (візуальна характеристика) та автоматична класифікація (контрольована та некерована класифікація). Найбільш перспективним є напрямок Барталева О.М. [7], роботи якого присвячені розробці методів оцінки стану і динаміці лісів на основі даних космічних спостережень. Особлива увага приділена методам контролю стану довкілля за допомогою систем дистанційного зондування, методам розпізнавання та обробці зображень, математичного моделювання, аналізу часових серій даних, математичній статистиці, геоінформатиці, системного та прикладного програмування. Просторово-часові методи досліджень лісових масивів на основі кореляційного синтезу ДЗЗ відображено в роботі Триснюка В.М. та Триснюка Т.В. [8]. Високі досягнення застосування ГС технологій відмічені в галузі сільського господарства. Прогнозування врожайності, проведення моніторингу полів для визначення загроз стало доступним завдяки використанню супутникових даних. Датчики видають зображення в різних спектрах, що дає можливість застосування чисельних спектральних індексів, таких як нормалізований диференційований вегетаційний індекс (NDVI), індекс вмісту хлорофілу в покривах (CCCI), що врегульовує внесення поживних речовин, нормалізований індекс Red Edge (NDRE) визначає вміст азоту, модифікований ґрунтово-керований вегетаційний індекс (NSAVI) застосовується для мінімізації впливу ґрунтового фону на ранніх фазах розвитку рослин [9]. Серед сучасних досліджень варто відмітити роботи Поморцевої О. Є. з моделювання розташування екологічно небезпечних об'єктів за допомогою ГС [10], Костюченко Ю.В., які в своїх працях зазначають високі перспективи використання ГС-технологій в вивченні навколишнього середовища [11], Бурштинської Х., Долинської І. [12], Кохан С.С. [13-14] та зарубіжних вчених Grégoire J-M., Pinnock S [15], Jones P.D., Moberg A [16], С. Donald Ahrens [17] та інших.

В процесі роботи над системою відслідковування змін рослинності було досліджено ряд комп'ютерних платформ для аналізу та візуалізації природних змін.

Веб-додаток GIOVANNI, розроблений GES DISC, забезпечує простий та інтуїтивний спосіб візуалізації, аналізу та доступу до величезної кількості даних дистанційного зондування Землі без необхідності завантаження даних. Система дозволяє отримати доступ до даних з декількох віддалених сайтів, підтримує кілька форматів даних, включаючи ієрархічний формат.

Вчені з Університету Колорадо в США запустили проект «Земний архів» (The Earth Archive). Його метою є спроба оцифрування ландшафту Землі з використанням «лазерних карт», отриманих за допомогою лідарів.

Аналіз сучасних систем для відстеження змін в глобальній екосистемі дійшов висновку, що всі вони дають недостатню кількість інформації щодо аналізу та візуалізації даних про рослинність. Більшість з них має ряд недоліків:

- повна або обмежена функціональність для аналізу та візуалізації місцевих даних;
- незважаючи на наявність інструкцій щодо використання цих інструментів, існує високий поріг входу для використання програми, що значно скорочує кількість користувачів;
- обмежені можливості бачення мікрокліматичних показників.

Таким чином, актуальність роботи полягає у відсутності великої кількості аналогів, які можуть бути використані для отримання достатньої інформації про поточний та попередній стан лісових насаджень і проведення порівняльного аналізу відповідних показників з метою прогнозування подальшого стану.

Метою дослідження є розробка інформаційної системи аналізу геоданих для відслідковування змін рослинності Чорнобильської зони із застосуванням нормалізованого вегетаційного індексу рослинності (NDVI) та методу кластеризації К-середніх.

Результати дослідження

Стрімкий розвиток сучасних засобів ДЗЗ значно розширив можливості отримання даних для аналізу змін рослинності на Землі. Передові технології дозволяють ефективно вирішувати завдання, пов'язані з отриманням якісно нової інформації, яка необхідна для сталого управління природною рослинністю та охороною природи.

Для вирішення поставлених завдань у дослідженні Чорнобильської зони була зібрана серія знімків супутникової програми Landsat 5 та 8 досліджуваної території за 2000-2020 роки. Для завантаження космічних знімків обрано відкритий електронний ресурс <http://earthexplorer.usgs.gov/>. Специфікація для вибору знімків наведена в табл. 1.

Таблиця 1 – Специфікація для вибору знімків

| Критерій | Необхідні дані |
|------------------------|---|
| Координати | м. Чорнобиль (51°16'24" п. ш., 30°13'20" с. д.) |
| Дата зйомки | 01.07.2000-08.31.2020 (липень-серпень) |
| Відсоток хмарності | Не більше 10 % |
| Рівень обробки | З проведеною геометричною корекцією |
| Формат отриманих даних | GeoTiff |

Слід зауважити, що під час проведення аналізу за допомогою супутникових зображень хмарність була мінімальною. Усі сцени, записані для аналізу, характеризують стан рослинності в липні-серпні кожного досліджуваного року, оскільки саме в літній період рослинність знаходиться у фазі вегетації. Для дослідження використовуються зображення із середньою просторовою роздільною здатністю, але площа сцени значно більша, ніж досліджувана область, тому

всі досліджені зображення будуть «приспосовані» до потрібної області з відповідним географічним посиленням для подальшої роботи. Також проводиться синтезація каналів для формування відповідного мультиспектрального зображення, радіометрична та атмосферна корекція знімків в програмі QGIS 3.

Для оцінки кількості рослинності на основі багатоспектральних космічних знімків використовуються різні індекси рослинності: NDVI (нормалізований диференційний індекс рослинності), WdVI (зважений диференціальний індекс вегетації), PVI (вертикальний індекс вегетації) [18]. В нашому дослідженні ми використовуємо нормалізований індекс NDVI, який характеризує густоту рослинного покриття. NDVI – нормалізований диференційний показник вегетації – кількісний показник фотосинтетичних властивостей активної біомаси, що застосовується для вирішення задачі кількісної оцінки рослинності, показників біологічної продуктивності сільськогосподарських культур та кормових угідь.

Обчислюється NDVI за формулою

$$NDVI = \frac{NIR - RED}{NIR + RED}, \quad (1)$$

де NIR – амплітуда відбиття в ближній інфрачервоній області спектра, RED – амплітуда відбиття в червоній області спектра.

Завдяки використанню при розрахунках NDVI нормалізованої різниці між мінімумом і максимумом відбиття забезпечується зниження різних форм мультиплікативного шуму (різниця в освітленні знімка, хмарності, серпанку, поглинання радіації атмосферою), наявного у спектральних каналах різночасових знімків. У проведеному дослідженні враховано особливості NDVI, пов'язану з помірною чутливістю до зміни ґрунтового фону (за винятком випадків, коли густина рослинного покриття була нижчою за 30 %).

Розрахунок NDVI базується на двох найбільш стабільних областях (незалежно від інших факторів) спектральної кривої відбиття судинних рослин. NDVI можна обчислити із зображень високої, середньої або низької роздільної здатності зі спектральними каналами в червоному діапазоні (0,55-0,75 мкм) та інфрачервоному (0,75-1,0 мкм).

Відповідно до формули (1.1), щільність рослинності (NDVI) у певний момент зображення дорівнює різниці між інтенсивністю відбитого світла в інфрачервоній та червоній зонах, поділеною на суму їх інтенсивності. Максимальне поглинання сонячної радіації хлорофілом вищих судинних рослин знаходиться в червоній зоні спектру (0,6-0,7 мкм), а область максимального відбиття структур клітин знаходиться в інфрачервоній зоні (0,7-1,0 мкм).

У даній роботі використовуються знімки супутників Landsat 5 з параметрами каналів:

3 (червоний, 0,63 – 0,69 мкм),

4 (ближній інфрачервоний, 0,76-0,90 мкм)

та Landsat 8 з параметрами каналів

4 (червоний, 0,64 – 0,67 мкм),

5 (ближній інфрачервоний, 0,85-0,88 мкм).

Під час роботи додатку обчислюється значення яскравості кожного пікселя вхідних зображень у видимому червоному та ближньому інфрачервоному каналах шляхом комбінації значень червоного, зеленого та синього кольору.

На рис. 1 зображено один і той самий фрагмент ділянки поверхні в різні роки (2000, 2010, 2020). Натуральний кольоровий композит використовує смугову комбінацію червоного, зеленого та синього. Він відтворюється близько до того, що бачить наше людське око. У той час як здорова рослинність зелена, нездорова – коричнева. Міські об'єкти виглядають білими та сірими, а вода – темно-синьою або чорною.

На рис. 2 зображені знімки в інфрачервоному спектрі. Ця комбінація смуг також називається композитом ближнього інфрачервоного. Він використовує ближній інфрачервоний та зелений. Оскільки хлорофіл відображається поблизу інфрачервоного світла, то ця смугова композиція корисна для аналізу рослинності. Зокрема, червоні ділянки мають кращий стан рослинності. Темні ділянки – вода, а міські – білі.

Через особливості відображення в різних областях червоного та інфрачервоного спектрів природні об'єкти, не пов'язані з рослинністю, мають постійне значення NDVI. У табл. 2 наведена шкала NDVI, в якій зазначається, що кожне значення NDVI відповідає різним об'єктам на землі.



Рис. 1. Знімки в «натуральних» кольорах Чорнобильської зони за 2000, 2010 та 2020 роки відповідно
(**Fig. 1.** Images in the natural colors of the Chernobyl zone for 2000, 2010 and 2020 respectively)

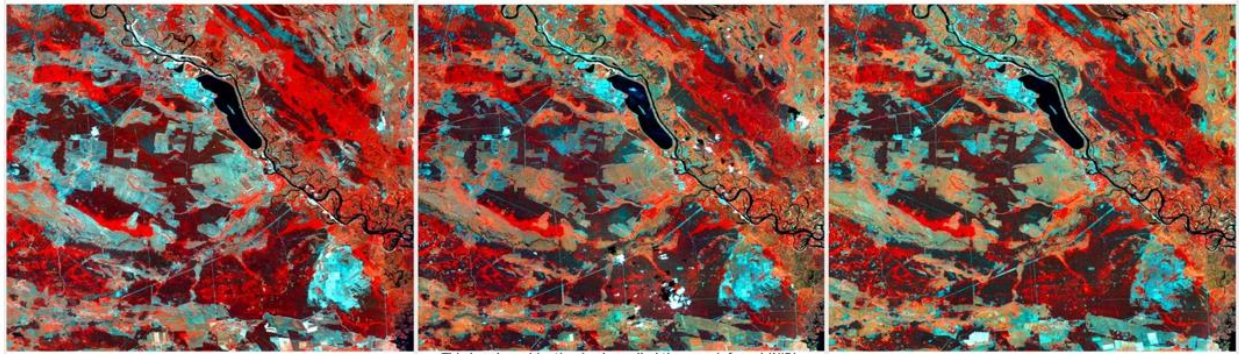


Рис. 2. Знімки в інфрачервоному спектрі Чорнобильської зони за 2000, 2010 та 2020 роки відповідно (Fig. 2. Images in the infrared spectrum of the Chernobyl zone for 2000, 2010 and 2020 respectively)

Таблиця 2 – Шкала NDVI

| Тип об'єкта | Відбиття в червоній області спектра | Відбиття в інфрачервоній області спектра | Значення NDVI |
|-----------------------------------|-------------------------------------|--|---------------|
| Густа рослинність | 0.1 | 0.5 | 0.7 |
| Розріджена рослинність | 0.1 | 0.3 | 0.5 |
| Відкритий ґрунт | 0.25 | 0.3 | 0.025 |
| Хмари | 0.25 | 0.25 | 0 |
| Сніг і лід | 0.375 | 0.35 | -0.05 |
| Вода | 0.02 | 0.01 | -0.25 |
| Штучні матеріали (бетон, асфальт) | 0.3 | 0.1 | -0.5 |

За цими значеннями обчислюється значення NDVI для кожного пікселя в межах від -1 до 1, рис. 3:

- -1.0-0 – відмерлі рослини, об'єкти неживої природи, вода;
- 0-0.16 – відкритий ґрунт;
- 0.17-0.40 – трав'яна рослинність;
- 0.41-0.60 – розріджена рослинність;
- 0.61-0.70 – густа рослинність;
- 0.71-1.0 – дуже густа рослинність.



Рис. 3. Дискретизована шкала NDVI (Fig. 3. Discretized scale NDVI)

Проблема відображення даних розв'язана через надання кожному відповідному пікселю вхідного зображення відповідного кольору, отриманого завдяки конвертації значення NDVI даного пікселя. Оцінюючи поверхню NDVI за 2020 рік, встановлено, що частка дуже густої рослинності набагато більша, ніж за попередні роки, відкритого ґрунту майже немає. На основі аналізу даних нормалізованого вегетаційного індексу за досліджений період побудовано карти змін щільності рослинного покриву за 2000, 2010, 2020 роки.

Результати моделювання, що приведені на рис. 4, наочно вказують на зміни, які відбулися на відповідних ділянках.

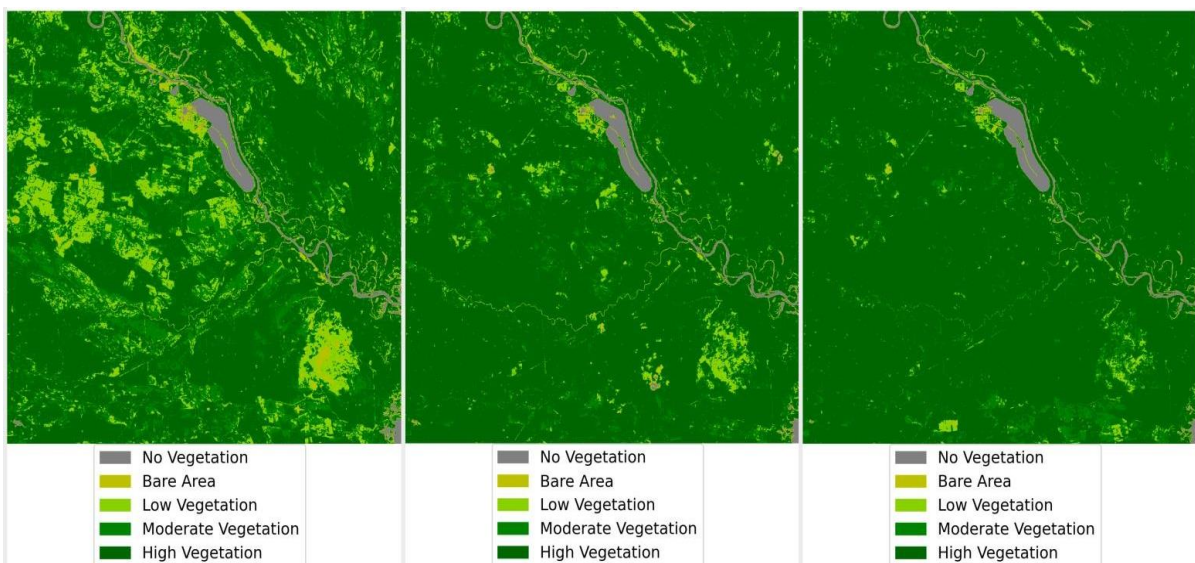


Рис. 4. Класифікація знімків Чорнобильської зони відчуження за розрахованими індексами вегетації за 2000, 2010 та 2020 роки відповідно (Fig. 4. Classification of images of the Chernobyl zone according to the calculated vegetation index according to 2000, 2010 and 2020 respectively)

Для визначення кількісних коливань у часі територій, на яких змінюється щільність рослинного покриття, пораховано кількість пікселів для кожного з розрахованих значень показника NDVI. За результатами класифікації на рис. 5 представлено діаграму кількісного порівняння рослинності.

Використання методу індексу рослинності дає змогу на підставі зміненого спектрального образу більш ефективно інтерпретувати об'єкти поверхні в Чорнобильській зоні.

Якість рослинного покриття за багатоспектральними космічними зображеннями оцінюється за вмістом важливих для життєдіяльності пігментів, та базується на аналізі спектральних характеристик відбиття живого листя рослин у видимому та ближньому інф-

рачервоному спектральних діапазонах. Оцінки, отримані за допомогою індексу NDVI, містять найбільш детальну інформацію про розподіл і різноманітність рослинного світу, визначення ділянок, на яких існують відхилення в рості і розвитку рослин. На рис. 6 показано проходження змін рослинного покриття за показниками вегетаційного індексу. Дані зображення свідчать, що в останні роки на території відчуження проходять суттєві зміни в рослинному покритті. Йде інтенсивне заростання лук та перелогів Чорнобильської зони деревною рослинністю (вже більше 50 %). Проте площі лісів поступово зменшуються. Це відбувається завдяки підтопленням, радіаційному ураженню, пожежам. Суттєвої шкоди лісу також завдають природні шкідники та хвороби.

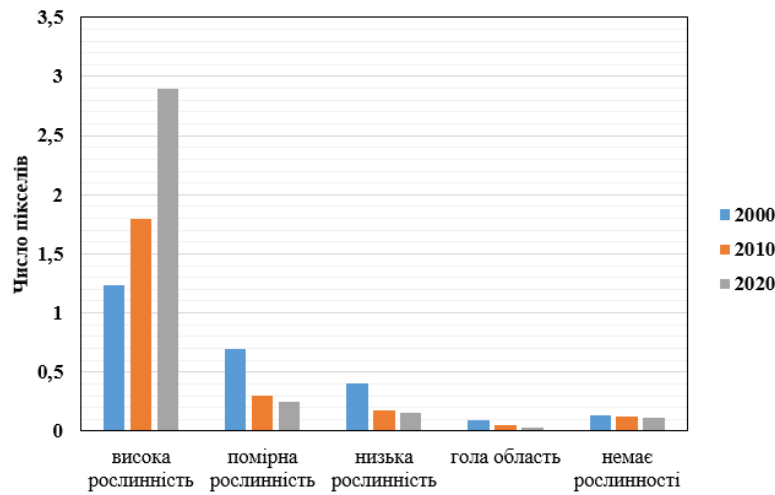


Рис. 5. Діаграма кількісного порівняння класів за розрахованим індексом вегетації за роками
(Fig. 5. Diagram of quantitative comparison of classes according to the calculated vegetation index by years)

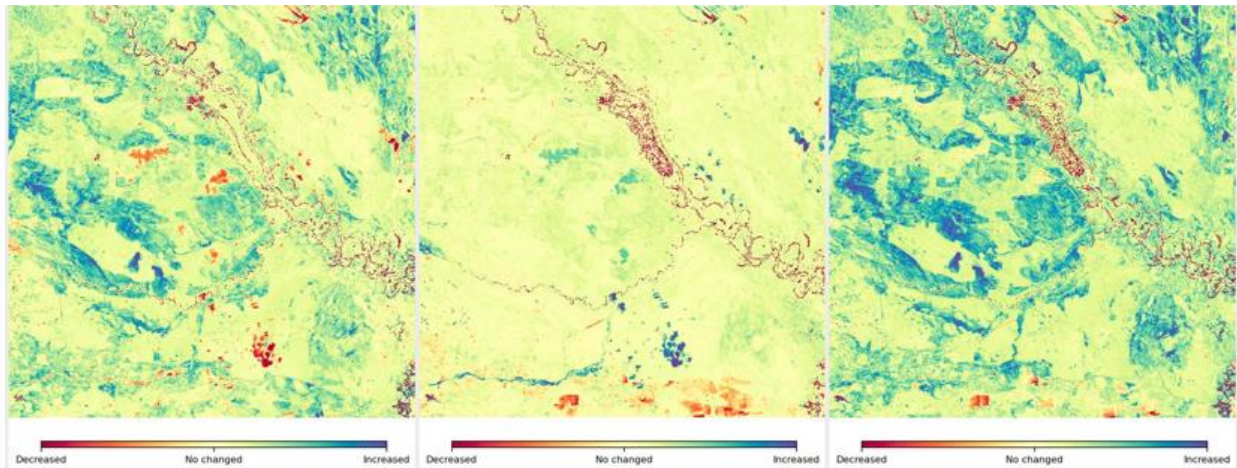


Рис. 6. Відображення змін рослинності Чорнобильської зони за 2000, 2010 та 2020 роки відповідно
(Fig. 6. Reflection of changes in the vegetation of the Chernobyl zone for 2000, 2010 and 2020 respectively)

Класифікація за даним індексом вегетації не несе вичерпної інформації про інші об'єкти на поверхні досліджуваної області, наприклад, вода, ліси. Також індекс NDVI у значній мірі залежить від кількості активної біомаси, тобто він стає неефективним в періоди, коли рослинність є послабленою або спостерігається низький рівень вегетації. У зв'язку з цим найбільш поширеним і використовуваним методом дешифрування космічних знімків є візуальне дешифрування знімка,

при якому розшифровку проводить експерт, який знає характеристики області та властивості об'єктів, представлених на зображенні [19]. Однак цей метод трудомісткий і досить тривалий, тому важливо дослідити методи автоматичного дешифрування (класифікації). Автоматична класифікація – це процес розбиття пікселів неперервного растрового зображення на категорії на основі їх спектральних значень, в результаті чого кожному пікселю присвоюється нове значення.

Оскільки апріорної інформації про об'єкт зйомки у нас немає, ми будемо використовувати *некеровану класифікацію* (кластеризація). Дана класифікація дозволяє виділити контури з неконтрастною структурою по спектральній яскравості – рослини, воду, хмари та інші об'єкти, розбивши зображення на групи пікселів, подібних за спектральними властивостями, тобто кластери. Зауважимо, що даний алгоритм потребує мінімум вхідних даних. Найбільш популярними методом кластеризації є алгоритми К-середніх [20], який розділяє n об'єктів на k кластерів, так щоб кожний об'єкт належав до кластера з найближчим до нього середнім значенням. За допомогою цього методу ми визначаємо породний склад дерев з наших знімків.

Кластеризація методом К-середніх (K-means) – це неконтрольований алгоритм машинного навчання, який використовується для поділу заданого набору D , що містить n об'єктів, на k різних кластерів C_1, C_2, \dots, C_k , що не перекриваються. Даний метод простий в реалізації та відносно швидкий. Об'єктивна функція використовується для оцінки якості розподілу, щоб об'єкти в кластері були схожими один на одного, але відрізнялися від об'єктів в інших кластерах. Це означає, що цільова функція спрямована на високу схожість у кластері та низьку схожість між кластерами.

Спочатку визначається кількість кластерів, що необхідно утворити. Потім набір даних D з n об'єктів розбивається на k кластерів так, щоб кожний об'єкт належав до кластеру з найближчим до нього середнім значенням c_i – центроїдом кластеру C_i .

Метод базується на мінімізації суми квадратів відстаней між кожним об'єктом та центром його кластера. Різниця між об'єктом p кластеру C_i і c_i вимірюється за допомогою метрики $d = dist(p, c_i)$, де $dist(x, y)$ – евклідова відстань між двома точками x і y . Цільова функція обчислюється за формулою:

$$E = \sum_{i=1}^k \sum_{p \in C_i} (p - c_i)^2, \quad (2)$$

де E – сума помилки у квадраті для всіх об'єктів у наборі даних, k – число кластерів, C_i – отримані кластери, p – точка в просторі, що представляє даний об'єкт, c_i – центроїд кластеру C_i .

Об'єкт до його кластерного центру розміщується в квадраті, а відстані підсумовуються. Ця цільова функція намагається зробити отримані k кластери максимально компактними та максимально відокремленими.

Алгоритм К-середніх:

1. Визначається кількість кластерів, що необхідно утворити.
2. Випадковим чином обирається k об'єктів, які на цьому кроці вважаються центрами кластерів.
3. Кожний об'єкт «приписується» до одного з n кластерів – того, відстань до якого найкоротша.
4. Розраховується новий центр кожного кластера як елемент, ознаки якого розраховуються як середнє арифметичне ознак об'єктів, що входять у цей кластер.
5. Ітерації продовжуються до того часу (повторюються попередні кроки 3-4), поки кластерні центри стануть стійкими, дисперсія всередині кластера буде мінімізована, а між кластерами – максимізована.

Обчислювальна складність алгоритму К-середніх дорівнює $O(nkt)$, де n – загальна кількість об'єктів, k – кількість кластерів, t – кількість ітерацій. Зазвичай $k \ll n$ і $t \ll n$. Тому метод є порівняно масштабованим та ефективним при обробці великих наборів даних.

На рис.7 зображено класифікацію знімків, за якими можна виявити змін в її структурі досліджуваної поверхні, а також для відокремлення рослинності від інших об'єктів та встановлення пікселів відкритого ґрунту.

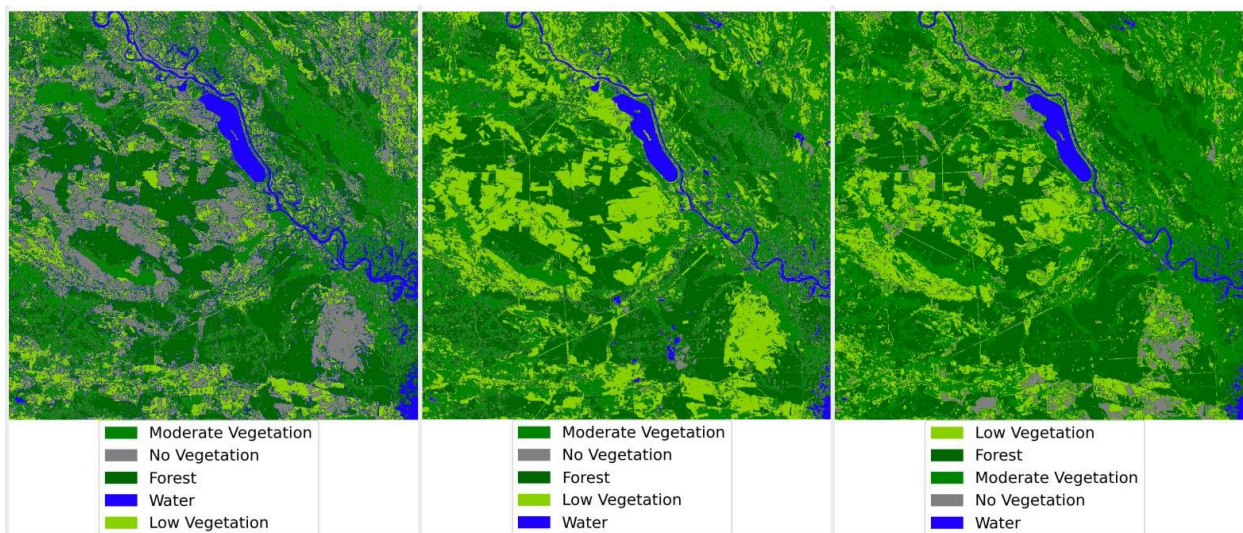


Рис. 7. Класифікація знімків Чорнобильської зони відчуження за допомогою методу К-середніх за 2000, 2010 та 2020 роки відповідно (Fig. 7. Classification of images of the Chernobyl zone using the method of K-means for 2000, 2010 and 2020 respectively)

За результатами класифікації на рис. 8 представлено діаграму кількісного порівняння рослинності за 2000, 2010, 2020 роки з використанням методу К-середніх.

Згідно результатів, наведених на рис. 7, 8, найбільшу частину у Чорнобильській зоні складають ліси, площа яких займає 62 % від усієї площі. На рис.8 видно, що з 2000 року по 2010 рік площа лісів

зросла приблизно на 15 %, а в період з 2010 року по 2020 роки навпаки трохи зменшилася – приблизно на 5 %. Це пояснюється тим, що в 2015 році та в 2020 році територією відчуження пройшли великі лісові пожежі, причинами яких стали висока температура, зменшення кількості опадів, сильні поривчасті вітри, і, відповідно, масове ураження дерев шкідниками й хворобами.

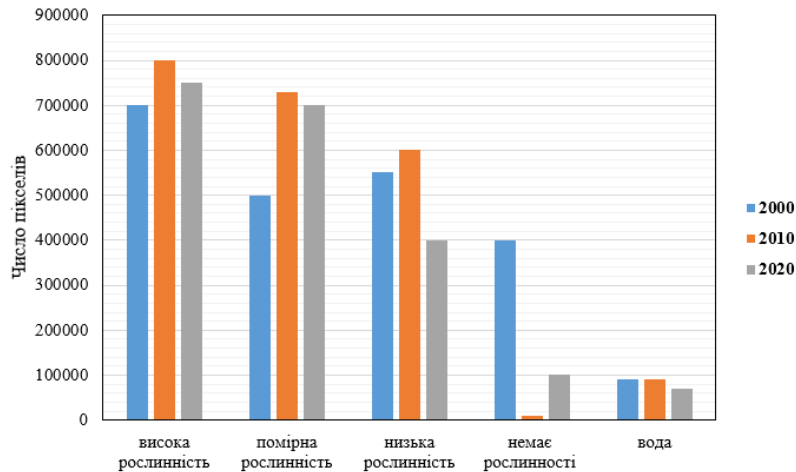


Рис. 8. Діаграма кількісного порівняння класів за 2000, 2010, 2020 роки з використанням методу К-середніх (Fig. 8. Diagram of quantitative comparison of classes for 2000, 2010 and 2020 years using the method of K-means)

Це все призводить до того, що ліс швидко відновитися не може, тому й кількісні показники заліснення території останніми роками впали. Чорнобильська зона відчуження має площу 260 тис. га. Наприклад, в 2015 році площа лісової пожежі досягла майже 400 га. В 2020 році внаслідок квітневої пожежі загалом згоріло 11500 га лісу, що становить 5 % всієї території, з них приблизно 35 % – ліси. Внаслідок інтенсивного горіння були втрачені або сильно пошкоджені певні рослинні групи. Проте, незважаючи на всі природні катаклізми (підвищення температури, засуха, зимові аномалії опадів та температур, буревії, лісові пожежі), а також людську діяльність (санітарні рубки, несанкціонована вирубка дерев), ліси продовжують рости і їх площі зростатимуть, хоча і не так швидко.

Перевіримо отримані результати про залісненість території Чорнобильської зони з використанням відомої формули Байєса:

$$P(B \setminus A) = \frac{P(A \setminus B) \cdot P(B)}{P(A)}, \quad (3)$$

де $P(B)$ – априорна ймовірність події B , $P(A)$ – ймовірність настання події A , $P(A \setminus B)$ – ймовірність настання події A при істинності гіпотези B (умовна ймовірність), $P(B \setminus A)$ – ймовірність гіпотези B за умови, що подія A вже настала (апостеріорна ймовірність).

Застосуємо формулу (3) для нашого випадку:

$$P(\text{ліс} \setminus \text{пож.}) = \frac{P(\text{пож.} \setminus \text{ліс}) \cdot P(\text{ліс})}{P(\text{пож.})} = \frac{P(\text{пожежі} \setminus \text{ліс}) \cdot P(\text{ліс})}{P(\text{пож.} \setminus \text{ліс}) \cdot P(\text{ліс}) + P(\text{пож.} \setminus \neg \text{ліс}) \cdot P(\neg \text{ліс})}, \quad (4)$$

де $P(\text{пож.} \setminus \text{ліс})$ – ймовірність настання пожежі в лісі, $P(\text{пож.} \setminus \neg \text{ліс})$ – ймовірність настання пожежі за межами лісу, $P(\neg \text{ліс})$ – ймовірність того, що на даній території лісу немає, $P(\text{ліс} \setminus \text{пож.})$ – ймовірність появи лісу при умові, що пожежа пройшла.

Підрахувавши площі територій, де були пожежі та лісові пожежі за період 2015-2020 років, ми за формулою (4) отримуємо, що заліснення території Чорнобильської зони має становити 70 % при ймовірності негоди 10 % за останні роки.

Порівнюючи отримані результати з тими, які ми мали б отримати теоретично, можемо зробити висновок, програмне забезпечення може давати похибку через некоректну класифікацію та невірні дані, адже ми не маємо можливості використати зображення 100 % якості, особливо, коли такі космічні знімки наразі платні.

Відповідно, якщо покращити алгоритм класифікації та працювати зі знімками вищої якості, то програма буде давати більш точніші результати.

Висновки

У статті розроблено інформаційну систему аналізу часових змін рослинності протягом десятиліть. Проаналізовано результати визначень нормалізованого вегетаційного індексу рослинності NDVI та за серією космічних знімків TM Landsat 5 та 8 протягом літнього періоду за 2000, 2010, 2020 роки для Чорнобильської зони відчуження.

Також встановлено ефективність використання методу К-середніх для оцінки зміни рослинності на даній території.

Для перевірки точності отриманих результатів використовувався метод Байєса, який довів нам, що дані методи є досить ефективними для аналізу динаміки стану рослинних угруповань в часі і дозволяють виявити загальні тенденції позитивної-негативної зміни рослинного покриву на досліджуваній території.

Даний програмний продукт є досить гнучким та універсальним, він може бути легко адаптованим для застосування не тільки для оцінки стану рослинного

покриву Чорнобильської зони, але і для інших територій із обмеженим доступом.

Подальшим напрямком вдосконалення розробленої інформаційної системи є впровадження штучного інтелекту, що надало б можливості більш ефективно та точно робити аналіз та оцінку класифікації, а також подальше вдосконалення й самостійна розробка алгоритму класифікації зображень без додаткових програмних застосунків.

СПИСОК ЛІТЕРАТУРИ (REFERENCES)

1. Michael N. DeMers. *Fundamental of geographic information systems*. New Mexico State University. 2010. 506 p.
2. Бутенко О.С., Барабаш О.В., Горелик С.І., Никітін А.А. Система космічного моніторингу за сміттєзвалищами твердих побутових відходів [Space monitoring system for landfill for operational detection of unauthorized landfills]. *Системи управління, навігації та зв'язку*. Випуск 2 (48). Полтава: ПНТУ, 2018. С. 114-119. DOI: <https://doi.org/10.26906/SUNZ.2018.2.114>.
3. Бутко І.М. Інформаційна технологія використання геопросторових інформаційних структур [Information technology of using geospatial information structures]. *Системи управління, навігації та зв'язку*. Випуск 2 (64). Полтава, ПНТУ, 2021. С. 4-10. DOI: <https://doi.org/10.26906/SUNZ.2021.2.004>.
4. Kuchuk G., Kovalenko A., Komari I.E., Svyrydov A., Kharchenko V. Improving Big Data Centers Energy Efficiency: Traffic Based Model and Method / Kharchenko V., Kondratenko Y., Kasprzyk J. (eds) *Green IT Engineering: Social, Business and Industrial Applications. Studies in Systems, Decision and Control*. Vol 171. Cham: Springer, 2019. DOI: https://doi.org/10.1007/978-3-030-00253-4_8
5. Коваленко А.А., Кучук Г.А. Методи синтезу інформаційної та технічної структур системи управління об'єктом критичного застосування [Methods for synthesis of informational and technical structures of critical application object's control system]. *Сучасні інформаційні системи [Advanced Information Systems]*. Т. 2, № 1. 2018. С. 22-27. DOI: <https://doi.org/10.20998/2522-9052.2018.1.04>
6. 30 років Чорнобильської катастрофи [30 years of the Chernobyl disaster]. Огляди. Збірник інформаційно-аналітичних доповідей. К.: КІМ, 2016. 170 с.
7. Bartalev S., Isaev A. Monitoring Northern Eurasian Terrestrial Ecosystems from Satellites: New Products and Perspective Approaches. *31st ISRSE Conference on Global Monitoring for Sustainability and Security*. St Petersburg, 2012.
8. Триснюк В.М., Охарев В.О., Триснюк Т.В. Створення системи мобільного екологічного моніторингу [Creating a system of mobile environmental monitoring]. *Екологічна безпека та збалансоване ресурсокористування*. № 2. 2018. С. 116-123.
9. IPCC Intergovernmental Panel on Climate Change: IPCC special report on land use, land use change and forestry. Cambridge University Press, 2012.
10. Поморцева О.Є. Моделювання розташування екологічно небезпечних об'єктів за допомогою геоінформаційних систем [Modeling the location of environmentally hazardous objects using geographic information systems]. *Вчені записки ТНУ імені В.І. Вернадського. Серія: технічні науки*. Т. 29(68), № 6, Ч.2. 2018. С. 222-226.
11. Костюченко Ю. В., Мовчан Д. М., Костюченко Ю. В., Артеменко І. Г., Білоус Ю. Г. Оцінка невизначеностей при визначенні продуктивності рослинного покриву та формуванні вуглецевого балансу територій за даними ДЗЗ [Estimation of uncertainties in determining the productivity of vegetation and the formation of the carbon balance of the territories according to remote sensing of the Earth]. *Геоінформатика*, № 3 (47). 2013. С. 67-75.
12. Буруштинська Х., Долинська І. Врахування впливу атмосфери під час опрацювання космічних зображень [Taking into account the influence of the atmosphere during the processing of space images]. *Геодезія, картографія і аерофотознімання*. Вип. 76. 2012. С. 70-73.
13. Кохан С.С. Дослідження динаміки вегетаційних індексів для оцінювання стану сільськогосподарських культур на основі даних IRS-1D LISS-III [Study of the dynamics of vegetation indices for assessing the condition of crops on the basis of data IRS-1D LISS-III]. *Вісник геодезії та картографії*. № 4(73). 2011. С. 20-24.
14. Кохан С.С. Застосування вегетаційних індексів на основі серії космічних знімків IRS-1D LISS-III для визначення стану посівів сільськогосподарських культур [Application of vegetation indices based on a series of space images IRS-1D LISS-III to determine the status of crops]. *Космічна наука і технологія*. Т. 17, № 5. 2011. С. 58.
15. Grégoire J.-M., Pinnock S. The World Fire Web network. A satellite based system for globally mapping fires in vegetation. *Publication of the European Communities*, S.P.I.00.11 2010. P. 6.
16. Jones P.D., Moberg A. Hemispheric and large-scale surface air temperature variations: an extensive revision and an update to 2001. *Journal of Climate*. 16. 2003. P. 206-223.
17. Donald C. Ahrens. *Meteorology Today*. Cengage Learning, 2016. 662 p.
18. Бардиш Б., Буруштинська Х. Використання вегетаційних індексів для ідентифікації об'єктів земної поверхні [Use of vegetation indices to identify objects on the earth's surface]. *Сучасні досягнення геодезичної науки та виробництва*. № 2 (28). 2014. С. 82-88.
19. Омелич І.Ю., Яременко А.А., Неповшвайленко Н.О., Горай І.В. Визначення тенденцій розвитку рослинного покриву на підставі розрахунку нормалізованого вегетаційного індексу на прикладі Петриківського району Дніпропетровської області [Determination of trends in the development of vegetation on the basis of the calculation of the normalized vegetation index on the example of Petrykivskiy district of Dnipropetrovsk region]. *Український журнал дистанційного зондування Землі*. Вип. 23, 2019. С. 9-13.
20. Coates A., Ng A. Y. *Learning Feature Representations with K-means*, Stanford University, 2012.

Received (Надійшла) 11.08.2021

Accepted for publication (Прийнята до друку) 20.10.2021

Барабаш Олег Володимирович – доктор технічних наук, професор, професор кафедри автоматизації проектування енергетичних процесів і систем, Національний технічний університет України «КПІ імені І. Сікорського», Київ, Україна;
Oleg Barabash – Doctor of Technical Sciences, Professor, Professor of the Department of Automation of projection of power processes and systems, National Technical University of Ukraine “Igor Sikorsky KPI”, Kyiv, Ukraine;
e-mail: bar64@ukr.net; ORCID ID: <https://orcid.org/0000-0003-1715-0761>.

Бандурка Олена Іванівна – аспірантка кафедри автоматизації проектування енергетичних процесів і систем, Національний технічний університет України «КПІ імені Ігоря Сікорського», Київ, Україна;
Olena Bandurka – Phd student of the Department of Automation of projection of power processes and systems, National Technical University of Ukraine “Igor Sikorsky KPI”, Kyiv, Ukraine;
e-mail: o.i.bandurka@ukr.net; ORCID ID: <https://orcid.org/0000-0002-8059-1861>.

Шпурик Вадим Вадимович – кандидат технічних наук, доцент кафедри автоматизації проектування енергетичних процесів і систем, Національний технічний університет України «КПІ імені Ігоря Сікорського», Київ, Україна;
Vadim Shpuryk – Candidate of Technical Sciences, Senior Lecturer of the Department of Automation of projection of power processes and systems, National Technical University of Ukraine “Igor Sikorsky KPI”, Kyiv, Ukraine;
e-mail: wadimoff@gmail.com; ORCID ID: <https://orcid.org/0000-0002-3477-5731>.

Свинчук Ольга Василівна – кандидат фізико-математичних наук, доцент кафедри автоматизації проектування енергетичних процесів і систем, Національний технічний університет України «КПІ імені Ігоря Сікорського», Київ, Україна;
Olha Svychnuk – Candidate of Physic-Mathematical Sciences, Senior Lecturer of the Department of Automation of projection of power processes and systems, National Technical University of Ukraine “Igor Sikorsky KPI”, Kyiv, Ukraine;
e-mail: 7011990@ukr.net; ORCID ID: <https://orcid.org/0000-0001-9032-6335>.

Информационная система анализа геоданных для исследования изменений растительности

О. В. Барабаш, Е. И. Бандурка, В. В. Шпурик, О. В. Свинчук

Аннотация. Быстрый рост возможностей геоинформационных технологий в области обработки и анализа пространственных данных привел к существенному росту роли геоинформационных систем в разных областях человеческой деятельности. Применение новых подходов к обработке пространственной информации со спутников для более эффективной и оперативной оценки состояния растительных покровов обусловлено растущей тенденцией доступности данных дистанционного зондирования Земли. В статье предложена информационная система, которая позволяет быстро и удобно отслеживать изменения растительного мира. Анализ проводился на примере Чернобыльской зоны за период 2000-2020 годов. Чернобыльская катастрофа совпала с периодом интенсивного вегетационного развития растений. В этот период они наиболее чувствительны к радиации. Было установлено, что наиболее часто для определения количественного состояния биомассы применяется индекс NDVI в разные временные интервалы. Но данный индекс становится неэффективным в периоды ослабления активной фазы вегетации. В этой связи представляется практический интерес для оценки возможности применения методом К-средних для анализа космических снимков растительного покрова на разных фазах вегетации. В результате исследования корректно интерпретирована водная поверхность, земли с растительным покровом и без него, благодаря чему определены земли с разреженной растительностью и плотным растительным покровом. Построены карты растительного покрова по нормализованному вегетационному индексу методом К-средних, по которому можно четко проследить динамику изменения растительности в течение 20 лет. Точность результатов была проверена методом Байеса. Согласно проведенным расчетам определено, что несмотря на все природные катаклизмы (повышение температуры, засуха, зимние аномалии осадков и температур, ураганы, лесные пожары), а также человеческую деятельность (санитарные рубки, несанкционированная вырубка деревьев), растительность в Чернобыльской зоне продолжает расти и ее площади будут расти, хотя и не так быстро.

Ключевые слова: дистанционное зондирование Земли; атмосферная коррекция космических снимков; Quantum GIS; нормализованный дифференциальный индекс растительности; алгоритм K-Means; растительность.

Information system of analysis of geodata for tracking changes of vegetation

Oleg Barabash, Olena Bandurka, Vadim Shpuryk, Olha Svychnuk

Abstract. The rapid growth of geo-information technology capabilities in the field of spatial data processing and analysis has led to a significant growth of the role of geo-information systems in different areas of human activity. Application of approaches to spatial information processing from satellites new for more effective and efficient assessment of the state of plant cover is caused by growing tendency of availability to data of Earth remote sensing. The article offers an information system that allows to quickly and conveniently track changes in the vegetation. The analysis was carried out on the example of the Chernobyl Area between 2000 and 2020. The Chernobyl Disaster coincides with the period of intensive vegetative plant development. During that period, they are most sensitive to radiation. It has been established that for defining the quantitative state of biomass the NDVI index at different time intervals is most often used. But this index becomes ineffective during periods of weakening of active phase of vegetation. This is therefore of practical interest to assess the possibility of using the K-means clustering for the analysis of space images of vegetation cover at different phases of vegetation. As a result of the research, water surface, land with and without vegetation has been correctly interpreted, thus determining the land with a sparse vegetation and dense vegetation cover. The maps of the vegetation cover according to the normalized vegetative index using the K-medium method were constructed, the method by which changes in vegetation over 20 years can be clearly observed. The accuracy results were verified with the Common Method Bias. According to the calculations, despite all natural cataclysms (temperature increase, drought, winter anomalies of precipitations and temperatures, storms, forest fires), as well as human activity (sanitary clear cuttings, illegal logging), vegetation in the Chernobyl zone continues to grow and its areas will increase, although not so quickly.

Keywords: Remote Earth sensing; atmospheric correction of space images; Quantum GIS; normalized differential vegetation index; K-means algorithm; vegetation.